A Multi-Method Exploration of Crime Hot Spots: SaTScan Results

by
Eric S. Jefferis[1]

Crime Mapping Research Center
National Institute of Justice
810 7th Street, NW
Washington, DC 20531
202.616.7108
jefferis@ojp.usdoj.gov

Paper presented at the
Annual Meeting of the *Academy of Criminal Justice Sciences (ACJS)*
Albuquerque, New Mexico

March 11, 1998

[1] Points of view are those of the author and do not necessarily represent the view of the U. S. Department of Justice or the National Institute of Justice. The author would like to thank Nancy La Vigne and the rest of the CMRC staff for their help in conceptualizing this project and writing this paper.

A Multi-Method Exploration of Crime Hot Spots: SaTScan Software Findings

Introduction

Many criminal justice agencies are taking advantage of rapid technological advancements in computer hardware and software in their efforts to more effectively and expeditiously identify unusual patterns of criminal activity within their jurisdictions. These unusual clusters (hereafter "hot spots") of crime across time and space are being identified and explored by various methods; from very simple visual interpretation of point data to the calculation of standard deviational ellipses to the animation of raster map images. Questions remain, though, about the appropriateness and validity of many hot spot identification techniques. It is generally thought that many of the hot spot identification methods have their individual strengths and weaknesses, but we are aware of no systematic comparison of the techniques to verify these assumptions. This paper is part of a "Multi-Method Exploration of Crime Hot Spots" that seeks to provide a systematic comparison of eleven hot spot identification techniques using common burglary and robbery data[2] provided by the Baltimore County Police Department.

Background

During September of 1997, the Crime Mapping Research Center at the National Institute of Justice held a meeting of various experts in the field of crime mapping to begin a dialogue on the issue of crime hot spot analysis and to establish a means to continue that dialogue. Specifically, the meeting was convened to: (1) establish what questions are left to be answered regarding hot spot analyses, (2) to begin to identify limitations of current methods, and (3) to brainstorm about possible new methods (if necessary). Based upon the discussion of these issues, the participants felt that an in-depth look at available techniques was a warranted and necessary first step[3]. As such, several of the participants agreed to participate in a research project to systematically compare eleven techniques of hot spot identification--many of which are currently used by law enforcement practitioners and researchers.

SaTScan Summary

SaTScan is a software program that has been designed to analyze spatial (and temporal) data with the *Spatial or Space-time Scan Statistic* (hereafter spatial scan statistic) developed by Kulldorff (1997; see also Kulldorff et al, 1997; Kulldorff and Nagarwalla, 1995). As the SaTScan documentation notes, the program was designed to: (1) evaluate spatial or space-time disease clusters to see if they are statistically significant, (2) test whether a disease is randomly distributed over space or over space and time, and (3) perform geographical surveillance of disease, to detect areas of significantly high or low rates. Though developed for epidemiological purposes, Kulldorff

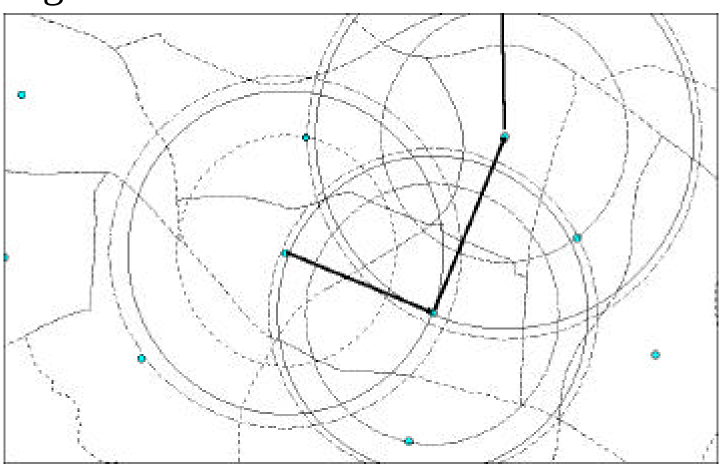
[2] We would like to extend our appreciation to Phil Canter of the Baltimore County Police Department for providing the data for this project. The Baltimore County data were drawn from the Regional Crime Analysis System (RCAS) for the period of November 1, 1996 through November 31, 1997. 7,719 records were included, of which 96.5% were geocoded.
[3] A list of participants and the meeting proceedings are available from the author upon request.

(1997:1482) notes that the spatial scan statistic may be applied to a variety of settings[4], so long as the analysis is conditional on the total number of observed points.

Spatial Scan Statistic

As applied to the identification of crime hot spots, the spatial scan statistic tests for the presence of statistically significant clusters as follows. SaTScan imposes a circular scan window centered on "each of several possible centroids[5] positioned throughout the study region (Kulldorff, 1997)". For each centroid, the window varies in size from zero to an upper limit imposed by the user. The upper limit indicates the maximum proportion of the population study area to be included in the window. Theoretically, this method results in an infinite number of distinct scan windows, with differing sets of neighbors within the circles, each of which is a possible cluster candidate. (The space-time scan statistic works in essentially the same fashion except that the circles become cylinders with their heights corresponding to time.) The maximum size of the scan window should be determined *a priori* rather than making numerous runs, varying the maximum of the scan window. Kulldorff warns against such repeated testing (to find which maximum size "best fits" the distribution) so as to maintain the statistical rigor of the method. Figure 1 provides an illustration of how a variably sized scan window might move across a point distribution.

Figure 1. Scan Window Movement



As can be inferred from the discussion above, the SaTScan program requires that the underlying population distribution be known. As such, point data must be aggregated to the same level of analysis as the population data. Fortunately, the program will accommodate data aggregated to all levels of analysis as long as the total population for that level of study is also available. For this analysis, data were aggregated to the census tract level; however, examining the data at block group, block, or housing unit level would also have been possible.

The spatial scan statistic operates on the assumption that the number of cases in each area is Poisson distributed. Under the null hypothesis of spatial randomness, the expected number of cases in each window is proportional to its population size. In this way, the spatial scan statistic adjusts for uneven population distributions.

Maximum Likelihood Ratio Function

For each scan window, a likelihood ratio test is conducted to test the hypothesis that there is an elevated rate of crime incidents in the scan window as compared to the distribution outside the

[4] Interestingly, Kulldorff (1997) does not mention crime analysis as a potential non-epidemiological application. Examples of applications he does provide include the spatial clustering of trees, celestial bodies, and uranium deposits.

[5] It is unclear at this point whether or not **all** centroids have scan windows imposed on them.

window. Assuming a Poisson distribution, the likelihood function for a specific window is proportional to:

$$(n/m)^n\,([N-n]/[N-m])^{(N-n)}\,I()$$

N is the total number of crimes across the entire study region and n is the number of cases in the scan window. m is the covariate-adjusted expected number of cases within the window under the null hypothesis, and $I()$ is an indicator function which is set to 1 if the window has more cases than are expected under the null hypothesis of spatial randomness. SaTScan maximizes the likelihood function over all windows identifying the window of the most likely cluster and secondary clusters, which are ranked according to their likelihood ratio. This likelihood ratio test allows SaTScan to scan for clusters with unusually low rates or for both areas of high and low rates.

SaTScan obtains the distribution and p-value of the most likely and secondary clusters by conducting Monte Carlo replications of the data set for each identified cluster. The user determines how many Monte Carlo replications are to be performed.

Ease of Use

SaTScan is a Windows 95 program, however, it has "not yet" been integrated into a Geographic Information System (Hjalmars, Kulldorff, Gustafsson, and Nagarwalla 1996). As such, several idiosyncrasies of the system should be noted. First, input data for SaTScan must be in space delimited ASCII format. Specifically, the user must create three ASCII input files—case file, population file, and coordinate file. The coordinate file has to be in integer format and this provided an added small complication. These requirements are more annoying and tedious than difficult. To illustrate, twelve input files (3 per study area) had to be prepared in order to conduct the four analyses presented here. Second, in addition to a results file, SaTScan outputs an ASCII "GIS File" that must be imported to a GIS and joined to the census tract map in order to graphically illustrate those areas identified as hot spots. (See Appendix B for an example of SaTScan's GIS output.) Finally, I found that the program inexplicably was overwriting output files when additional runs were performed.

Despite these complications, only approximately 2 ½ hours had passed from the time this first-time-user began to download SaTScan to the time when I had produced the four maps for this project.

A final positive note, the help files and on-line documentation provided with the program are outstanding. The developer has organized the "Contents" of the help file in such a manner that a forty-five page manual can be printed for easy desk reference.

Face Validity

One of the greatest strengths of this program is that it identifies hot spots in consideration of the underlying population. Though they appear reasonable and valid, it is difficult to subjectively look at an incident point map and infer whether the hot spots that the program identified are appropriate. I should note also at this point that SaTScan is capable of simultaneously identifying areas of both high and low concentrations, across time, and in consideration of multiple covariates. Obviously, it would be very difficult and time consuming to visually examine a series of maps in order to identify unusual clustering of events across time. Considering multiple covariates, I dare say, would make the task impossible. As variables are added, face validity very rapidly becomes difficult to assess.

Practical Utility

As noted above, SaTScan has not yet been integrated with a GIS package. As such, it probably has limited utility for most criminal justice practitioners. I would expect that very few line police officers would be interested in using this program on a regular basis to identify neighborhood hot spots of criminal activity. The processes involved in identifying hot spots with SaTScan would be time prohibitive. Many more crime analysts and researchers, though, would find the program useful, particularly in that it provides measures of statistical significance.

Flexibility

The program allows the user to set several parameters, including the number of Monte Carlo replications to perform for each identified cluster and the maximum size of the search window. Recall from above that the search window varies in size from zero up to coverage of a pre-defined proportion of the population. The developer recommends a maximum window size of 50% of the population. SaTScan can be set to search for high or low rates of incidents or both simultaneously.

Recall also from the previous discussion that SaTScan is capable of identifying clusters across time and also in consideration of categorical covariates. These properties alone make the program very attractive to both researchers and crime analysts.

Finally, an interesting option that deserves mention, but that I have yet to explore, is SaTScan's ability to accommodate three-dimensional data.

SaTScan Analysis of Baltimore County Robberies and Burglaries

For the demonstration purposes of this project, a purely spatial analysis was conducted with the maximum geographical cluster size set at 20% of the total population and the number of Monte Carlo replications set at 999[6]. Table 1 presents summary information on the "most likely" clusters for each of the four study areas. For a complete listing of results, readers are encouraged to refer to the SaTScan output presented in Appendix A. The discussion that follows focuses solely on the single "most likely" cluster as identified by SaTScan for each area of study.

Table 1. "Most Likely" Clusters of Burglaries and Robberies

	Burglaries	*Robberies*	*SW Burglaries*	*SW Robberies*
Study Area				
# of census tracts	199	199	67	67
Total population	692,134	692,134	236,181	236,181
Total cases	6,054	1,188	2,065	497
Annual cases/100,000	46.0	9.0	46.0	11.1
"Most likely" cluster				
# of census tracts*	2	30	1	16
Total population	4,140	100,012	131	45,563
Total cases	157	342	16	174
Annual cases/100,000	199.6	18.0	642.9	20.1
Expected # of cases	36.2	171.7	1.2	95.9
Log likelihood ratio	110.7	80.5	27.4	33.7
Monte Carlo rank	1/1000	1/1000	1/1000	1/1000
p-value	.001	.001	.001	.001

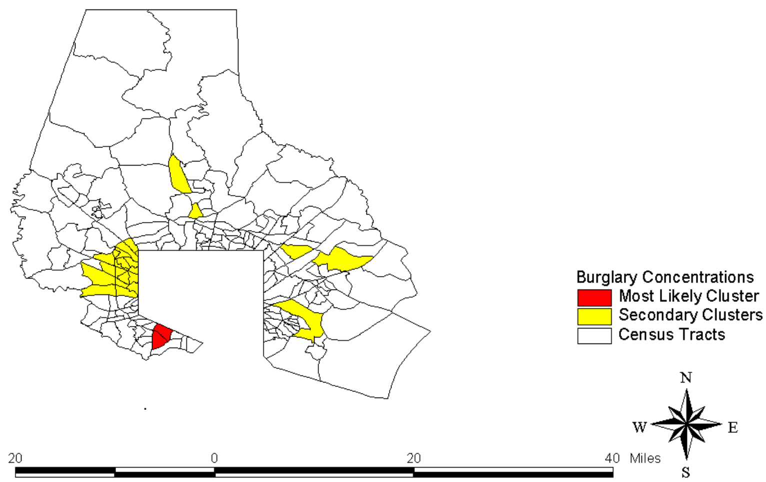
*A listing of census tracts included in each cluster can be found in the SaTScan output presented in Appendix A.

At the county level, 6054 burglaries were observed during the

[6] Setting the number of Monte Carlo replications at 9999 resulted in the same hot spots being identified, but at a p-level of .0001.

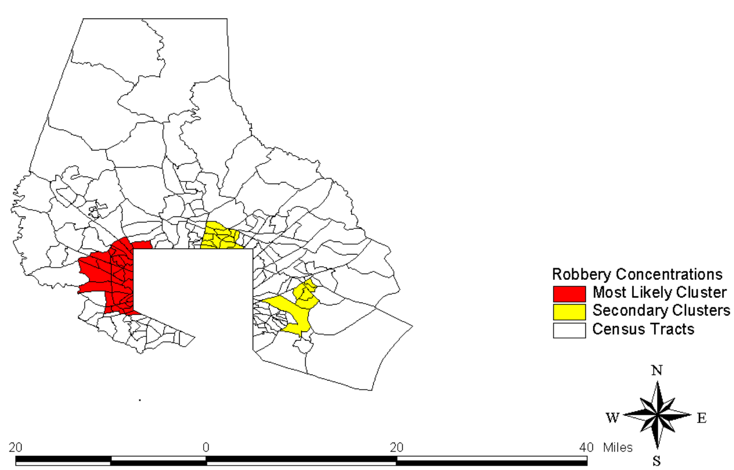
period of study, a rate of 46 offenses per 100,000 population. The most likely cluster identified in this region contained two census tracts and had a log likelihood ratio of 110.7 (p<.001). For this cluster, 36.2 cases were expected and 157 observed. The Monte Carlo rank for this cluster is 1/1000. Figure 2 depicts the census tracts identified as the most likely and secondary clusters of burglaries at the county level.

Burglary Concentrations

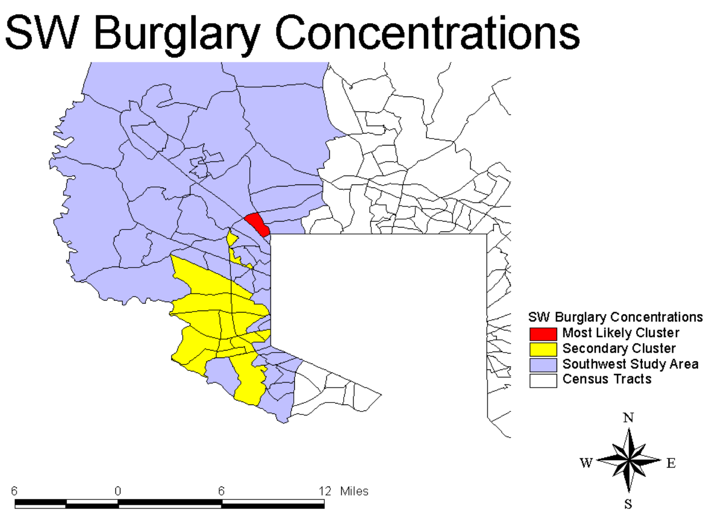


Significantly fewer robberies (n=1188) than burglaries were observed at the county level—or 9.0 per 100,000 population. The most likely robbery cluster was much larger than was the most likely burglary cluster, consisting of 30 census tracts. For this cluster, 171.7 cases were expected and 342 were observed. A log likelihood ratio was 80.5. Figure 3 depicts the census tracts identified as the most likely and secondary clusters of robberies at the county level.

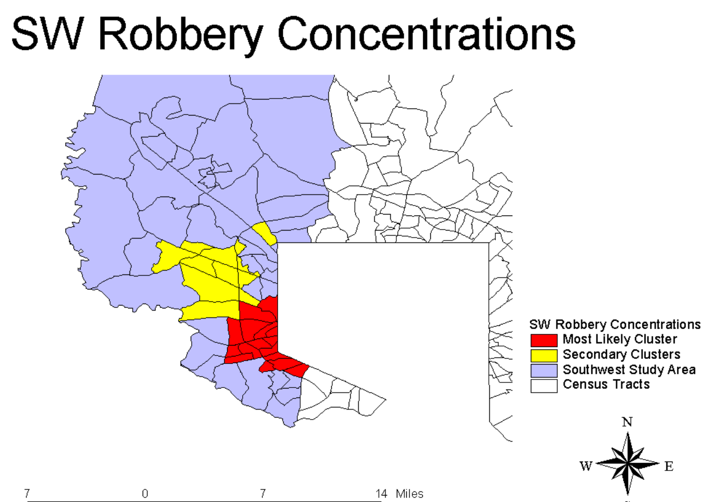
Robbery Concentrations



For burglaries in the southwest study region, their were 2065, or 46 per 100,000 population. SaTScan identified one census tract as being the most likely cluster. This area had a reported sixteen burglaries for a population of only 131. The spatial scan statistic expected only 1.2 cases in this region (log likelihood ratio of 27.4). This raises an interesting question about the accuracy and appropriateness of using population statistics for an analysis of burglary offenses. Had housing units been used to control for the area's underlying distribution of at-risk places, the results may have differed. Figure 4 depicts the census tracts identified as the most likely and secondary clusters of burglaries in the southwest study region.



Finally, 497 robberies occurred in the southwest region (11.1 per 100,000 population) and SaTScan identified a hot spot consisting of 16 census tracts. Within the "most likely" robbery cluster, 95.9 robbery cases were expected and 174 observed. The log likelihood ratio for this cluster was 33.7. It is interesting to note that, at both the county-wide and southwest region areas of study, the most likely clusters for robberies consisted of significantly more census tracts than the respective burglary clusters. Figure 5 depicts the census tracts identified as the most likely and secondary clusters of robberies in the southwest study region.



Conclusion

Overall, SaTScan holds great promise as an additional crime analysis tool. Too often, hot spots are identified based solely on their spatial location with no consideration of underlying population (or housing units, number of retail establishments, etc) levels. This "denominator problem" is addressed by the spatial scan statistic. Caution should be exercised, though, in the selection of the proper denominator. For example, I have used census tract populations as the denominator in this study, but for burglaries at least, households may be more appropriate.

My primary recommendation to the developers would be to fully integrate the program with a GIS. Comments by the developer would seem to indicate that they are in the process of doing just that. A fully integrated program would eliminate the need to create the various ASCII input files and would enable the user to produce maps of significant clusters without the need for the joining process. Further, this will enable statistical results from the cluster output files to be linked to their respective graphical illustrations.

Finally, I would recommend that the developer consider enabling the user to test for cluster specific and global spatial autocorrelation. It may well be argued that the simple identification of significant clusters does not require correction for spatial autocorrelation, but for advanced users who may want to estimate multivariate models, this does become an issue.

Readers are encouraged to explore the utility of SaTScan for themselves. The program can be downloaded free of charge from: http://www.dcpc.nci.nih.gov/bb/Software.html

REFERENCES

Hjalmars, U., Kulldorff, M., Gustafsson, G., and N. Nagarwalla (1996) "Childhood Leukemaemia in Sweden: Using GIS and a Spatial Scan Statistic for Cluster Detection." Statistics in Medicine, 15:707-715.

Kulldorff, M. (1997) "A Spatial Scan Statistic." Communications in Statistics: Theory and Methods, 26:1481-1496.

Kulldorff, M. and N. Nagarwalla (1995) "Spatial Disease Clusters: Detection and Inference." Statistics in Medicine, 14:799-810.

Kulldorff, M., Feuer, E., Miller, B., and L. Freedman (1997) "Breast Cancer in Northeastern United States: A Geographical Analysis." American Journal of Epidemiology, 146:161-170.

APPENDIX A

“All Burglaries” Results:

```
                 ______________________________

                     SaTScan, version 1.0.1
                 ______________________________


Program run on: Wed Feb 18 11:11:57 1998
Purely Spatial analysis scanning for clusters with high rates.
______________________________________________________________________

SUMMARY OF DATA

Study period .........: 1973/1/1 - 1991/12/31
Number of census areas: 199
Total population .....: 692134
Total cases ..........: 6054
Annual cases / 100,000: 46.0
______________________________________________________________________

MOST LIKELY CLUSTER

Census areas included.: 4304, 4305
Coordinates / radius .: (-7.66812e+06,3.92417e+06) / 1204.14
Population ...........: 4140
Number of cases ......: 157            (36.21 expected)
Annual cases / 100,000: 199.6            (Relative risk: 4.34)
Log likelihood ratio..: 110.728668
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


SECONDARY CLUSTERS

Census areas included.: 4406
Coordinates / radius .: (-7.64811e+06,3.93696e+06) /  0.00
Population ...........: 967
Number of cases ......: 60            (8.46 expected)
Annual cases / 100,000: 326.6            (Relative risk: 7.09)
Log likelihood ratio..: 66.230940
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4524
Coordinates / radius .: (-7.64583e+06,3.92698e+06) /  0.00
Population ...........: 2302
Number of cases ......: 76            (20.14 expected)
Annual cases / 100,000: 173.8            (Relative risk: 3.77)
Log likelihood ratio..: 45.342458
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001
```

```
Census areas included.: 4084
Coordinates / radius .: (-7.66508e+06,3.94838e+06) /  0.00
Population ...........: 363
Number of cases ......: 31            (3.18 expected)
Annual cases / 100,000: 449.5            (Relative risk: 9.76)
Log likelihood ratio..: 42.877173
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4023.04, 4024.04, 4031, 4024.05, 4032.02,
                        4024.03, 4023.03, 4032.01, 4023.05, 4033,
                        4011.02, 4012, 4034.02, 4023.01, 4024.01,
                        4034.04, 4034.01, 4015.01
Coordinates / radius .: (-7.67371e+06,3.93405e+06) / 4295.89
Population ...........: 62176
Number of cases ......: 730            (543.84 expected)
Annual cases / 100,000: 61.8            (Relative risk: 1.34)
Log likelihood ratio..: 31.923197
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4517.01
Coordinates / radius .: (-7.64151e+06,3.93598e+06) /  0.00
Population ...........: 1790
Number of cases ......: 49            (15.66 expected)
Annual cases / 100,000: 144.1            (Relative risk: 3.13)
Log likelihood ratio..: 22.653689
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4088
Coordinates / radius .: (-7.66284e+06,3.94289e+06) /  0.00
Population ...........: 2316
Number of cases ......: 53            (20.26 expected)
Annual cases / 100,000: 120.5            (Relative risk: 2.62)
Log likelihood ratio..: 18.319768
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4411.01, 4409.02, 4501
Coordinates / radius .: (-7.65057e+06,3.93311e+06) / 1388.28
Population ...........: 8809
Number of cases ......: 129            (77.05 expected)
Annual cases / 100,000: 77.1            (Relative risk: 1.67)
Log likelihood ratio..: 14.756835
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4111.02, 4113.02
Coordinates / radius .: (-7.64052e+06,3.94382e+06) / 4777.43
Population ...........: 4840
```

```
Number of cases ......: 67            (42.33 expected)
Annual cases / 100,000: 72.9            (Relative risk: 1.58)
Log likelihood ratio..: 6.144037
Monte Carlo Rank .....: 202/1000
P-value ..............: 0.202


Census areas included.: 4903.02, 4909, 4912.01, 4912.02, 4908
Coordinates / radius .: (-7.65856e+06,3.94068e+06) / 2134.01
Population ...........: 14639
Number of cases ......: 161            (128.05 expected)
Annual cases / 100,000: 57.9            (Relative risk: 1.26)
Log likelihood ratio..: 4.009363
Monte Carlo Rank .....: 763/1000
P-value ..............: 0.763


Census areas included.: 4085.03
Coordinates / radius .: (-7.6634e+06,3.94649e+06) /  0.00
Population ...........: 3237
Number of cases ......: 43            (28.31 expected)
Annual cases / 100,000: 69.9            (Relative risk: 1.52)
Log likelihood ratio..: 3.299406
Monte Carlo Rank .....: 929/1000
P-value ..............: 0.929


Census areas included.: 4904
Coordinates / radius .: (-7.66556e+06,3.93972e+06) /  0.00
Population ...........: 1631
Number of cases ......: 24            (14.27 expected)
Annual cases / 100,000: 77.5            (Relative risk: 1.68)
Log likelihood ratio..: 2.757953
Monte Carlo Rank .....: 988/1000
P-value ..............: 0.988
```

```
The log likelihood ratio value required for an observed
cluster to be significant at level
... 0.01: 9.775882
... 0.05: 7.701261
```

```
Warning: According to the input data, the following tracts have a
         population totaling zero for the specified year(s).

         Tract 2604.98, 1990
         Tract 2606.99, 1990
         Tract 2805, 1990
         Tract 4522, 1990
```

PARAMETER SETTINGS

```
Analysis Type           ([1=Purely Spatial], 2=Space-Time)  : 1
Areas to Scan           ([1=High], 2=Low, 3=High and Low)   : 1

Cases File             : C:\Hotspot\ctburg.txt
Population File        : C:\Hotspot\ctpop.txt
Coordinates File       : C:\Hotspot\ctcoord.txt
Output File            : C:\Hotspot\all burgs.txt
Precision of Case Times (0=None, [1=Year], 2=Month, 3=Day)  : 0

Spatial Dimensions      ([2=2 Dimensions], 3=3 Dimensions)  : 2
Special Grid            ([0=No], 1=Yes)                     : 0
Max Geographic Size     (0-[50])                            : 20

Start Date              (YYYY, YYYY/MM, YYYY/MM/DD)   : 1973/1/1
End Date                (YYYY, YYYY/MM, YYYY/MM/DD)   : 1991/12/31

Replications            (9 or [999]-29999, ending in 999)   : 999
```

“**All Robberies” Results:**

```
        SaTScan, version 1.0.1

Program run on: Wed Feb 18 11:16:24 1998
Purely Spatial analysis scanning for clusters with high rates.
______________________________________________________________________

SUMMARY OF DATA

Study period .........: 1973/1/1 - 1991/12/31
Number of census areas: 199
Total population .....: 692134
Total cases ..........: 1188
Annual cases / 100,000: 9.0
______________________________________________________________________

MOST LIKELY CLUSTER

Census areas included.: 4024.03, 4032.02, 4024.04, 4012, 4024.05,
                        4023.04, 4032.01, 4011.02, 4031, 4034.04,
                        4013.01, 4023.05, 4013.02, 4034.02,
                        4011.01, 4033, 4023.03, 4007.02, 4015.05,
                        4034.01, 4007.01, 4035, 4023.01, 4024.01,
                        4006, 4015.01, 4008, 4002, 4009, 4023.02
Coordinates / radius .: (-7.67192e+06,3.93327e+06) / 6307.74
Population ...........: 100012
Number of cases ......: 342            (171.66 expected)
Annual cases / 100,000: 18.0            (Relative risk: 1.99)
Log likelihood ratio..: 80.541127
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


SECONDARY CLUSTERS

Census areas included.: 4511, 4505.02, 4505.01, 4508.02, 4524,
                        4508.01
Coordinates / radius .: (-7.6448e+06,3.9291e+06) / 2960.41
Population ...........: 24407
Number of cases ......: 110            (41.89 expected)
Annual cases / 100,000: 23.7            (Relative risk: 2.63)
Log likelihood ratio..: 40.147845
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4912.02, 4915, 4912.01, 4914, 4913, 4916,
                        4903.02, 4920.01, 4909, 4921.02, 4908,
                        4910, 4911
Coordinates / radius .: (-7.65752e+06,3.93892e+06) / 2560.25
Population ...........: 42276
```

```
Number of cases ......: 147            (72.56 expected)
Annual cases / 100,000: 18.3            (Relative risk: 2.03)
Log likelihood ratio..: 31.881676
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001


Census areas included.: 4409.02, 4411.01, 4405, 4404, 4410, 4408,
                        4411.02, 4501, 4512
Coordinates / radius .: (-7.65033e+06,3.93419e+06) / 2842.03
Population ...........: 32812
Number of cases ......: 85            (56.32 expected)
Annual cases / 100,000: 13.6            (Relative risk: 1.51)
Log likelihood ratio..: 6.672913
Monte Carlo Rank .....: 116/1000
P-value ..............: 0.116


Census areas included.: 2606.99, 4210, 4211.02, 4209, 4211.01
Coordinates / radius .: (-7.65365e+06,3.92507e+06) / 2067.23
Population ...........: 11381
Number of cases ......: 37            (19.53 expected)
Annual cases / 100,000: 17.1            (Relative risk: 1.89)
Log likelihood ratio..: 6.298732
Monte Carlo Rank .....: 150/1000
P-value ..............: 0.150


Census areas included.: 4406
Coordinates / radius .: (-7.64811e+06,3.93696e+06) /  0.00
Population ...........: 967
Number of cases ......: 5            (1.66 expected)
Annual cases / 100,000: 27.2            (Relative risk: 3.01)
Log likelihood ratio..: 2.178234
Monte Carlo Rank .....: 999/1000
P-value ..............: 0.999
```

```
The log likelihood ratio value required for an observed
cluster to be significant at level
... 0.01: 9.349251
... 0.05: 7.581824
```

```
Warning: According to the input data, the following tracts have a
         population totaling zero for the specified year(s).

         Tract 2604.98, 1990
         Tract 2606.99, 1990
         Tract 2805, 1990
         Tract 4522, 1990
```

PARAMETER SETTINGS

```
Analysis Type           ([1=Purely Spatial], 2=Space-Time)  : 1
Areas to Scan           ([1=High], 2=Low, 3=High and Low)   : 1

Cases File             : C:\Hotspot\ctrobb.txt
Population File        : C:\Hotspot\ctpop.txt
Coordinates File       : C:\Hotspot\ctcoord.txt
Output File            : C:\Hotspot\all robs.txt
Precision of Case Times (0=None, [1=Year], 2=Month, 3=Day)  : 0

Spatial Dimensions      ([2=2 Dimensions], 3=3 Dimensions)  : 2
Special Grid            ([0=No], 1=Yes)                     : 0
Max Geographic Size     (0-[50])                            : 20

Start Date              (YYYY, YYYY/MM, YYYY/MM/DD)   : 1973/1/1
End Date                (YYYY, YYYY/MM, YYYY/MM/DD)   : 1991/12/31

Replications            (9 or [999]-29999, ending in 999)   : 999
```

“Southwest Burglaries” Results:

```
                 _____________________________

                    SaTScan, version 1.0.4
                 _____________________________


Program run on: Wed Feb 18 14:54:01 1998
Purely Spatial analysis scanning for clusters with high rates.
______________________________________________________________________

SUMMARY OF DATA

Study period .........: 1973/1/1 - 1991/12/31
Number of census areas: 67
Total population .....: 236181
Total cases ..........: 2065
Annual cases / 100,000: 46.0
______________________________________________________________________

MOST LIKELY CLUSTER

Census areas included.: 4034.01
Coordinates / radius .: (-7.67231e+06,3.93793e+06) /  0.00
Population ...........: 131
Number of cases ......: 16            (1.15 expected)
Annual cases / 100,000: 642.9            (Relative risk: 13.97)
Log likelihood ratio..: 27.388713
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001

SECONDARY CLUSTERS

Census areas included.: 4015.04, 4015.03, 4015.01, 4015.05, 4009,
                        4014, 4010, 4024.01, 4008, 4011.01,
                        4007.01, 4024.05, 4011.02
Coordinates / radius .: (-7.67745e+06,3.92958e+06) / 4904.13
Population ...........: 46930
Number of cases ......: 540            (410.32 expected)
Annual cases / 100,000: 60.6            (Relative risk: 1.32)
Log likelihood ratio..: 23.839894
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001

Census areas included.: 4023.05
Coordinates / radius .: (-7.6745e+06,3.9359e+06) /  0.00
Population ...........: 1873
Number of cases ......: 47            (16.38 expected)
Annual cases / 100,000: 132.1            (Relative risk: 2.87)
Log likelihood ratio..: 19.159228
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001

Census areas included.: 4004.01
```

```
Coordinates / radius .: (-7.67301e+06,3.92505e+06) /  0.00
Population ...........: 4437
Number of cases ......: 68            (38.79 expected)
Annual cases / 100,000: 80.7            (Relative risk: 1.75)
Log likelihood ratio..: 9.169974
Monte Carlo Rank .....: 3/1000
P-value ..............: 0.003

Census areas included.: 4007.02
Coordinates / radius .: (-7.67183e+06,3.9287e+06) /  0.00
Population ...........: 1775
Number of cases ......: 32            (15.52 expected)
Annual cases / 100,000: 94.9            (Relative risk: 2.06)
Log likelihood ratio..: 6.742539
Monte Carlo Rank .....: 54/1000
P-value ..............: 0.054
_____________________________________________________________________

The log likelihood ratio value required for an observed
cluster to be significant at level
... 0.01: 7.747753
... 0.05: 6.777665


_____________________________________________________________________
PARAMETER SETTINGS

Analysis Type          ([1=Purely Spatial], 2=Space-Time)  : 1
Areas to Scan          ([1=High], 2=Low, 3=High and Low)   : 1

Cases File            : C:\Hotspot\subburg.txt
Population File       : C:\Hotspot\subpop.txt
Coordinates File      : C:\Hotspot\subcoord.txt
Output File           : C:\Hotspot\sub robs.txt
Precision of Case Times (0=None, [1=Year], 2=Month, 3=Day)  : 0

Spatial Dimensions     ([2=2 Dimensions], 3=3 Dimensions)  : 2
Special Grid           ([0=No], 1=Yes)                     : 0
Max Geographic Size    (0-[50])                            : 20

Start Date             (YYYY, YYYY/MM, YYYY/MM/DD)   : 1973/1/1
End Date               (YYYY, YYYY/MM, YYYY/MM/DD)   : 1991/12/31

Replications           (9 or [999]-29999, ending in 999)   : 999
_____________________________________________________________________
```

“Southwest Robberies” Results:

```
                  ________________________________

                     SaTScan, version 1.0.4
                  ________________________________


Program run on: Wed Feb 18 14:50:28 1998
Purely Spatial analysis scanning for clusters with high rates.
_____________________________________________________________________

SUMMARY OF DATA

Study period .........: 1973/1/1 - 1991/12/31
Number of census areas: 67
Total population .....: 236181
Total cases ..........: 497
Annual cases / 100,000: 11.1
_____________________________________________________________________

MOST LIKELY CLUSTER

Census areas included.: 4007.02, 4006, 4013.02, 4007.01, 4013.01,
                        4002, 4011.01, 4008, 4016.01, 4001,
                        4011.02, 4012, 4010, 4015.05, 4009, 4309
Coordinates / radius .: (-7.67183e+06,3.9287e+06) / 3390.11
Population ...........: 45563
Number of cases ......: 174            (95.88 expected)
Annual cases / 100,000: 20.1            (Relative risk: 1.81)
Log likelihood ratio..: 33.733143
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001

SECONDARY CLUSTERS

Census areas included.: 4024.01, 4023.03, 4023.02, 4023.01,
                        4015.01, 4026.02, 4023.05, 4023.04,
                        4025.03, 4024.05
Coordinates / radius .: (-7.67703e+06,3.9338e+06) / 3714.07
Population ...........: 44093
Number of cases ......: 161            (92.79 expected)
Annual cases / 100,000: 19.2            (Relative risk: 1.74)
Log likelihood ratio..: 26.624767
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001

Census areas included.: 4034.01
Coordinates / radius .: (-7.67231e+06,3.93793e+06) /  0.00
Population ...........: 131
Number of cases ......: 8            (0.28 expected)
Annual cases / 100,000: 321.4            (Relative risk: 29.02)
Log likelihood ratio..: 19.280102
Monte Carlo Rank .....: 1/1000
P-value ..............: 0.001
_____________________________________________________________________
```

```
The log likelihood ratio value required for an observed
cluster to be significant at level
... 0.01: 7.922877
... 0.05: 6.126044
```

```
PARAMETER SETTINGS

Analysis Type          ([1=Purely Spatial], 2=Space-Time)  : 1
Areas to Scan          ([1=High], 2=Low, 3=High and Low)   : 1

Cases File            : C:\Hotspot\subrob.txt
Population File       : C:\Hotspot\subpop.txt
Coordinates File      : C:\Hotspot\subcoord.txt
Output File           : C:\Hotspot\sub robs.txt
Precision of Case Times (0=None, [1=Year], 2=Month, 3=Day)  : 0

Spatial Dimensions     ([2=2 Dimensions], 3=3 Dimensions)  : 2
Special Grid           ([0=No], 1=Yes)                     : 0
Max Geographic Size    (0-[50])                            : 20

Start Date             (YYYY, YYYY/MM, YYYY/MM/DD)   : 1973/1/1
End Date               (YYYY, YYYY/MM, YYYY/MM/DD)   : 1991/12/31

Replications           (9 or [999]-29999, ending in 999)   : 999
```
